A STUDY OF THE CONSISTENCY IN NEW YORK STATE FIRST YEAR MATH EXAMS

Stephanie Schaefer
American College of Education
stephanie.schaefer@ace.edu

Deborah Moore-Russo
University of Oklahoma
dmr@ou.edu

As standards documents have been introduced over the past 20 years, many states have seen an evolution in both the standards and related high stakes exams. For many teachers across the U.S., the rollout of standards and exams has not been an experience that builds trust in state education leaders. In this study, we consider three major changes in the first-year high school math exams in New York State since 2002, looking at consistency in item types, topics addressed, and student performance. Shifts in all were noted, but the changes in topics, especially when not obvious by the names given to standards, are suggested as the mostly likely to misinform or misguide teachers. We consider how state educational leaders are working to build trust for the next iteration of standards. While this study is particular to one state, the methods and findings should be of interest to others who study curriculum and testing in high schools.

Keywords: Algebra and Algebraic Thinking; Assessment; High School Education; Standards

The National Council of Teachers of Mathematics (NCTM, 1989) introduced its first standards document just over 20 years ago. This launched an evolution of mathematics curriculum standards and related assessments. NCTM (2000) published a subsequent set of standards in just over 10 years, and the Common Core Standards were published 10 years later (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). While having a national organization, such as NCTM, put forth standards helped bring more coherence to mathematics instruction across the United States, this led to significant disruptions at the state level. State education leaders subsequently made changes in their standards and state assessments that were rolled out to teachers, while legislation (e.g., No Child Left Behind, 2002) put in place mandates that implemented state assessment accountability.

Researchers report that state assessments have been associated with intense stress for teachers, and teachers report feeling that the high-stakes state exams often undermine meaningful learning and instruction (Barksdale-Ladd & Thomas, 2000). This feeling of pressure is regardless of the state in which teachers work (von der Embse, Pendergast, Segool, Saeki & Ryan, 2016), and studies have reported on state assessment accountability and its impact on teacher stress and teachers' intentions to leave the teaching profession (e.g., Ryan et al., 2017).

A reconstitution of the mathematics curriculum based on standards has brought about research that compares states' curricula and adoption (Senk & Thompson, 2020) as well as research that considers states' suggested coverage and placement of key topics in the curricula (e.g., Nagle & Moore-Russo, 2014; Stanton & Moore-Russo, 2012). However, little has been published about the high-stakes state exams themselves. How much did changing standards impact state assessments, particularly those that were being administered for the first year of high school mathematics? In order to consider this question in light of a new set of state mathematics standards that will soon impact state assessments, the overarching research goal of this study is to look at changes in first-year high school mathematics state exams in New York State (NYS) over the past 20 years. The three specific research questions include:

1. How consistent have state exams been over the past 20 years, in terms of the topics covered, types of items included, and student performance on the exams?
2. How do previous exam topics compare to current first year mathematics state standards?
3. How are state education leaders preparing teachers for the upcoming changes in standards and state exams?

Theoretical Framework

Social capital theory can focus on notions ranging from power (Fine, 2001) to economic transactions (Hardin, 1999) to mechanisms that build common values (McNiell, 2007). In all of these, there is interconnectivity that involves some level of trust. *Trust* is the belief that an entity (be it an institution or an individual) will act in ways that are consistent with one's expectations of positive behavior (Algan, 2018). It is an expectation that arises that is based on shared norms (Fukuyama, 1996). Trust helps reduce uncertainty (Luhmann, 1979) and helps deliver optimal outcomes by strengthening relationships and preventing defection (Six, van Zimmeren, Popa & Frison, 2015). Hardin (2002) suggests that norms produce trust, which he describes in the first person as the perception that you have "an interest to take my interests in the relevant matter seriously…you value the continuation of our relationship" (p. 2).

In organizations, subordinates' trust in those over them has been found to be related to employee commitment and job satisfaction (Colquitt, Scott, & LePine, 2007; Dirks & Ferrin, 2002). Burke, Sims, Lazzara, and Salas (2007) suggest that this trust is based on aspects such as accountability, transparency, and consistency (Kim & Lee, 2018). Consider the mathematics education community in a state as an organized entity with teachers in roles that might be considered subordinate to those who create high-stakes exams. The state's educational leaders show trust in teachers with their administration and marking of the exams. However, if one has listened to teacher lounge grumblings or has read media reports (Domanico, 2021; Strauss, 2012; Taylor, 2016), it calls to question how much this trust has been reciprocated. To begin to understand teacher trust in state education leaders and mandated state testing, there are many aspects that could be studied. A logical initial investigation is to study how consistent state exams have been looking at different measures of consistency. For that reason, this study considers looks the three iterations of the NYS exams for first-year high school math in terms of consistency and using the current math standards as a lens. It also considers how NYS education leaders are preparing teachers for future revisions to the standards and related exams.

Current Study

Since this study focuses on a particular state, we first detail the sequential shifts that have occurred in first-year high school mathematics in NYS. Next, we situate ourselves and experiences as researchers for this study.

Context.

From 2002 to 2020, there were three changes in standards that affected the first-year math exams in NYS: the *old* standards (Math A) from 2002-2009, the *recent* standards (Integrated Algebra) from 2008-2015, and the *current* standards from 2014-2020. Each exam administration occurred in January, June, and August. Each exam consisted of multiple choice and constructed response items corresponding to the relevant standards. In some years, two versions of the exam were available during transitional periods between standards. There were a total of 2287 items that were graded and publicly available on the NYS Regents' website (NYS Office of State Assessment, 2021): 764 *old* items (2002-2009), 857 *recent* items (2008-2015), and 666 *current* items (2014-2020).

Research team.

The research team consists of two members who have firsthand experience with the changes in NYS standards and assessments in high school mathematics. The lead researcher is a former NYS secondary mathematics teacher who experienced: the *old* exams as a student, the *recent* exams as a pre-service teacher, and the *current* exams as an in-service teacher. As part of her dissertation work (Schaefer, 2020), Schaefer did an in-depth study of NYS state assessments focusing on readability. While focusing on readability, other changes across the years became obvious. The second researcher worked in a Department of Learning and Instruction in a university located in NYS; she taught the secondary mathematics education classes for pre-service teachers from 2004 to 2014, a period that overlapped the *old*, *recent* and *current* standards.

Data Collection and Standard Classification

All items from first-year *old* (n=764), *recent* (n=857), and *current* (n=666) NYS exams were used as data. A sample of 10% of *current* exam items were classified using the *current* standards to obtain interrater reliability with NYS standard classification (Schaefer, 2020). Cohen's κ was calculated to determine agreement ($\kappa = 0.903$, $p < 0.05$). Discrepancies were then analyzed, and the NYS classification was accepted in each area of discrepancy. All *old* and *recent* items were then coded by the first using the *current* standards. Both authors discussed any items that were difficult to code until consensus was reached.

Findings

In this section we present findings in the order of the research questions. The findings are then followed by relevant discussion.

Exam Consistency

We first consider how consistent the first-year mathematics exams have been in NYS. We look at the topics covered by the exams in light of the standards that applied at the time of the exams. We then consider the types of items on the exams. Finally, we investigate any changes in student performance.

Item Topics. The *old*, *recent*, and *current* exams had a wide variety of topics based on their individual standard systems. Table 1 notes the different topic areas for each of the exams in terms of the distributions by the relevant standards at the time. The *recent* standards show the emphasis on an integrated curriculum, with the inclusion of a Geometry standard, that was not present in either the *old* or *current* standards. The *old* exams had the greatest difference in topic names as compared to *recent* and *current* exams. Consider the *current* standard of Number and Quantity. On *old* exams, the Number and Numeration standard was the focus for 8% of exam items, and the Operations standard was the focus for 19% of exam items. On *recent* exams, Number Sense and Operations was the focus of 8% of exam items. On *current* exams, Number and Quantity was the focus of 5% of exam items. Measurement decreased from being the focus on 19% of items on the *old* exam to 6% on *recent* exams, and it was not considered a high school mathematics standard and was therefore not the focus of any items on the *current* exam. The emphasis on functions shifted from accounting for 18% of the items on *old* exams, under the Patterns/Functions standard, to being melded in with Algebra in *recent* exams. Functions is a standard by itself and represented the focus of 37% of the items on *current* exams.

Table 1: Topic Areas and Relative Distributions of Exams.

Old Exam		*Recent* Exam		*Current* Exam	
Standard	Dist.	Standard	Dist.	Standard	Dist.
Math Reasoning	7%	Algebra	54%	Algebra	50%
Operations	19%	Geometry	16%	Functions	37%
Number & Numeration	8%	Number Sense & Ops	8%	Number & Quant	5%
Model/Multiple Reps	20%	Stats & Probability	16%	Stats & Probability	9%
Measurement	19%	Measurement	6%		
Uncertainty	8%				
Patterns/Functions	18%				

Item Types. Exams since 2004 but prior to *current* exams seemed to be more heavily weighted toward multiple-choice items. When the *old* exams were initially administered, the relative weighting (in terms of point distribution) of multiple-choice items was 47%. Then, the *old* exam was revised so that the relative weighting of multiple-choice items was 71% of exam points. *Recent* exams had a similar 69% relative weighting of multiple-choice items. The *current* exams relative weighting decreased to 56% of the total points. Most notably, there was also a six-point constructed response item added to *current* exams, replacing the three-point items that were on previous exams. Table 2 outlines the differences in item type distributions with the relative weighting for each item type in relation to the overall total exam point value.

Table 2: Item Type Distributions and Relative Weighting to Total Point Values for Exams

Item Type & Value	*Old* Exam Jun02-Jun03		*Old* Exam Jan04-Jun09		*Recent* Exam Jun08-Jun15		*Current* Exam Jun14-Jan20	
	n	Rel Wt*	*n*	Rel Wt	*n*	Rel Wt	*n*	Rel Wt
Multiple-Choice (MC Items)								
2-point	20	47%	30	71%	30	69%	24	56%
Constructed Response (CR) Items								
2-point	5	12%	5	12%	3	7%	8	19%
3-point	5	18%	2	7%	3	10%	0	0%
4-point	5	24%	2	10%	3	14%	4	19%
6-point	0	0%	0	0%	0	0%	1	7%

Student Performance. In order to consider exam consistency, we now look at student performance on the exams. For the NYS Regents, passing an exam is equivalent to receiving at least 65% of the total points on an exam. However, some students with disabilities are given a 55% passing rate, depending on their Individualized Education Programs (IEPs). A score at or above 85% is considered passing with distinction. For this reason, Table 3 displays the passing rates at or above 65% and 85% for all students taking the first-year math exams in NYS as well as the passing rates at or above 55%, 65%, and 85% for students with documented disabilities who took these exams (NYS Education Department, 2022).

Note the passing rate of scores that are 65% or higher on exams has decreased for all students going from 73% on *old* exams to 72% on *recent* exams to 70% on *current* exams. The passing with distinction rates (i.e., scores of 85% or higher) have decreased more markedly for all students going from 25% on *old* exams to 16% on *recent* exams to 13% on *current* exams. Now consider only the population of NYS students with disabilities and their performance on the *old*, *recent*, and *current* exams. Students with disabilities' passing rates of 55% or higher have increased from 64% to 66% to 67%, respectively, while passing their passing rates of 65% have decreased from 45% to 42% to 39%, respectively. There has also been a decline in student with disabilities who pass with distinction (i.e., with scores of 85% or higher) from 6% to 2% to 1% respectively.

Table 3: Relative Passing Rates on Exams for All Students and Students with Disabilities

Item Type & Value	*Old* Exam n = 1,552,177			*Recent* Exam n = 1,933,213			*Current* Exam n = 1,324,731		
	≥55%	≥65%	≥85%	≥55%	≥65%	≥85%	≥55%	≥65%	≥85%
All Students n = 4,980,809	-	73%	25%	-	72%	16%	-	70%	13%
Students with Disabilities n = 700,048	64%	45%	6%	66%	42%	2%	67%	39%	1%

Comparison to *Current* Standards

We now use the *current* standards as a lens to consider shifts in the topics covered on exams for the past 20 years. When considering the items on *old*, *recent*, and *current* exams (as displayed in Table 4), it is obvious that there have been shifts in the topics that are covered. This is most notable when looking that the number of items that were on the *old* exams (73%) and *recent* exams (57%) that do not apply to the *current* standards. For example, measurement and geometry are no longer part of the first-year mathematics curriculum in NYS. For the Algebra standard, there has been a respective shift from 22% to 29% to 50% respectively on *old*, *recent*, and *current* exams, with the items on the *current* exams primarily focusing on this standard. For the Functions standard, there has also been a notable increase from 1% to 4% to 37% respectively on *old*, *recent*, and *current* exams. This is due to introducing the concept of function as well as an emphasis on introducing and using functional notation (rather than only considering equations, such as $y = 2x + 3$). There have been less dramatic shifts in the percentages of items that address the Number and Quantity and the Statistics and Probability standards. There have

also been some shifts within individual topics as to which subtopics are emphasized. However, most of these shifts are based on the large increases in all the subtopics under the Algebra and Function standards.

Table 4: Counts of Items on Exams by Current Standards' Topics/Subtopics

	Old		*Recent*		*Current*	
Topic and Subtopic Areas	*n*	%	*n*	%	*n*	%
Algebra	**166**	**22%**	**248**	**29%**	**331**	**50%**
Arith w/ Polynomials & Ratl Expressions	19	2%	21	2%	34	5%
Creating Equations	22	3%	38	4%	82	12%
Reasoning w/ Equations & Inequalities	109	14%	139	16%	144	22%
Seeing Structure in Expressions	16	2%	50	6%	71	11%
Functions	**7**	**1%**	**34**	**4%**	**244**	**37%**
Building Functions	-	-	1	< 1%	37	6%
Interpreting Functions	7	1%	25	3%	153	23%
Linear, Quadratic, & Exponential Models	-	-	8	1%	54	8%
Number and Quantity	**19**	**2%**	**27**	**3%**	**33**	**5%**
The Real Number System	2	< 1%	-	-	18	3%
Quantities	17	2%	27	3%	15	2%
Statistics and Probability	**13**	**2%**	**58**	**7%**	**58**	**9%**
Interpreting Categorical & Quantitative Data	13	2%	58	7%	58	9%
None Applicable	**559**	**73%**	**490**	**57%**	**-**	**-**

Preparations for Upcoming Changes in Standards

Another change in NYS is coming, with the Next Generation Math Standards (NYS Education Department, 2017a). When the *old* curriculum evolved to the *recent* curriculum, there was a significant shift in topic areas covered in the standards. In the shift from the *recent* to the *current* curriculum, similar topic areas were used; yet standards were condensed and there was a move away from an integrated curriculum. Now, from the *current* to the *future* Next Generation Math Standards, NYS educational leaders have created resources (i.e., a snapshot document and a crosswalk) to better explain how standards have been added, modified, or removed.

In the Snapshot document (NYS Education Department, 2019a), there were three categories that highlighted the major changes from *current* to *future* standards. These categories including mapping the *future* standards to the *future* curriculum, outlining which standards moved to

another level of the mathematics curriculum, and additional clarification on *current* standards pertinent to the *future* curriculum. The Crosswalk document (NYS Education Department, 2019b) included detail clarifying the standards based on input from conference calls hosted by the NYS Education Department. These comments included notes that clarified more information about the standards. For instance, there was specific information regarding fluency expectations and how the first-year standards differed from those for the third-year math course, Algebra II.

The NYS Education Department created two committees that included both educators and parents for the *future* standards. There were also opportunities for public commentary on the draft document with over 750,000 comments from more than 10,500 people on the AIMHighNY 2016 survey (NYS Education Department, 2017b). This information was first used to create an initial version of the future standards in 2017 that has since been revised and disseminated in June 2019.

Probably most importantly, there has also been a longer roll-out period for the *future* standards than existed in past years when transitioning between standards. Initial information on the *future* standards was provided to teachers in 2017. The NYS Department of Education has also considered the impact of the pandemic and are taking its impact into account regarding the implementation of the *future* standards, which are not slated to be implemented and used for Algebra 1 exams until September 2023 (NYS Education Department, 2021).

Discussion

In conclusion, there have been some shifts in first-year high school math exams in NYS over the past 20 years. This is noted both in the item types and their weightings. It is also noted in student performance on the exams, especially in students passing with distinction. However, the areas where the revisions have been the most dramatic, especially in the impact on teachers, are in the topics that are covered on these exams.

Much of the discontent with the roll out of standards in the past is likely due to the lack of communication and a disruption in the continuity of topics that were included in the standards and on the exams. Change, by its very nature, disrupts consistency. However, there are mechanisms that can make change easier. There are also practices that can downplay change so that it is less evident initially, but this has consequences when it is realized that the change was more dramatic than initially believed. Having standards with similar topic names helped teachers transition from *recent* to *current* standards. However, what counted as Algebra for the *recent* exams was not the same as what counted as Algebra in the *current* exams. While 54% of the *recent* exam items were classified as addressing Algebra topics under the *recent* standards; only 29% of the *recent* exam items were classified as Algebra topics under the *current* standards. For example, ratios and proportions (including the three basic right triangle trigonometric ratios) were considered part of the subtopics covered in the *recent* Algebra standards, but they are not part of the *current* Algebra standards. Using the same topic name for a standard that represents different subtopics misinforms and can misguide teachers. When Algebra means one thing for a few years and then something else, it could have felt like a breach of trust for many teachers.

It is heartening that the NYS Department of Education has been more transparent and open to suggestions from teachers (and others) in the transition to future math standards that will be adopted soon. One way that this is noted in the inclusion of accessible documents that carefully map out the continuity across and evolution in topics. Optimistically, the research team is hopeful that the rollout of the future standards will be smoother than it was for past transitions of standards in NYS.

The input of teachers (as well as of other constituents in the learning process) is key to building trust. But bi-directional communication between teachers and state education leaders is not all that is needed to help teachers transition between curricular standards. There needs to be a logical, coherent evolution that builds on, rather than totally replaces, existing standards. Moreover, this evolution needs to be clearly communicated to all, but especially to teachers. Consistent transparency as to the coming changes and how they are line up with, build on, and improve the current standards would be of benefit and would help develop more trust between teachers and state education leaders.

Limitations, Future Study

While an obvious limitation is that this study only looks at NYS, the methods used and the underlying message are pertinent to those who study high school mathematics in other states, especially those interested in curricular standards and state exams. Other studies that investigate how different states have handled curricular shifts in standards and state assessment exams would be of great value. It would be interesting to see how future work could springboard from studies such as this to delve into more qualitative work on how trust is built or breached between teachers and state education leaders, especially in light of high stakes testing.

References

Algan Y. (2018). Trust and social capital. In J. E. Stiglitz, J.-P. Fitoussi, & M. Durand (Eds.), *For good measure advancing research on well-being metrics beyond GDP* (pp. 283-320). Paris: OECD (Organisation for Economic Co-operation and Development).

Barksdale-Ladd, M. A., & Thomas, K. F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. *Journal of Teacher Education, 51*(5), 384-397.

Burke, C. S., Sims, D. E., Lazzara, E. H., & Salas, E. (2007). Trust in leadership: A multilevel review and integration. *The Leadership Quarterly, 18*(6), 606–632.

Colquitt, J. A., Scott, B. A., & LePine, J. F. (2007). Trust, trustworthiness, and trust propensity: Meta-analytic test of their unique relationship with risk-taking and job performance. *Journal of Applied Psychology, 92*(4), 909–927.

Domanico, R. (2021, September 13). *Reset N.Y. state tests to recognize reality.* New York Daily News. http://nydailynews.com/

Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: Meta-analysis findings and implications for research and practice. *Journal of Applied Psychology, 87(*4), 611–628.

Fine, B. (2001). *Social capital versus social theory: Political economy and social science at the turn of the millennium.* London: Routledge.

Kim, H., & Lee, T. H. (2018). Strategic CSR communication: A moderating role of transparency in trust building. *International Journal of Strategic Communication, 12(*2), 107-124.

Fukuyama, F. (1996). *Trust: The social virtues and the creation of prosperity.* New York: The Free Press.

Hardin, R. (1999). Social capital. In J. E. Alt & M. Levi (Eds.). *Competition and cooperation: Conversations with Nobelists about economics and political science.* New York: Russel Sage Foundation.

Hardin, R. (2002). *Trust and trustworthiness.* New York: Russel Sage Foundation.

Luhmann, N. (1979). *Trust and power*. Chichester: John Wiley & Sons.

McNiell, D. (2007). Social capital or sociality? Methodological contrast between economics and other social sciences. In S. Ioannides & K. Nielsen (Eds.). *Economics and the social sciences: Boundaries, interaction and integration* (pp. 163-184). North Hampton: Edward Elgar.

Nagle, C., & Moore-Russo, D. (2014). Slope across the curriculum: Principles and Standards for School Mathematics and Common Core State Standards. *The Mathematics Educator, 23*(2), 40-59.

National Council for Teachers of Mathematics. (1989). *Commission on Standards for School Mathematics. Curriculum and evaluation standards for school mathematics.* Reston, VA: NCTM.

National Council of Teachers of Mathematics. (2000). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: NCTM.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics.* Washington, DC: Authors.

New York State Education Department. (2017a). *New York State Next Generation Mathematics Learning Standards.* http://www.nysed.gov/common/nysed/files/programs/curriculum-instruction/nys-next-generation-mathematics-p-12-standards.pdf

New York State Education Department. (2017b). *Board of Regents P-12 Committee Approves Next Generation Learning Standards.* http://www.nysed.gov/news/2017/board-regents-p-12-committee-approves-next-generation-learning-standards

New York State Education Department. (2019a). *New York State Next Generation Mathematics Learning Standards Algebra 1 Snapshot.* http://www.nysed.gov/common/nysed/files/programs/curriculum-instruction/nys-math-standards-algebra-i-snapshot.pdf

New York State Education Department. (2019b). *New York State Next Generation Mathematics Learning Standards Algebra 1 Crosswalk.* http://www.nysed.gov/common/nysed/files/programs/curriculum-instruction/nys-math-standards-algebra-i-crosswalk.pdf

New York State Education Department. (2021). *New York State Next Generation Mathematics Learning Standards Instruction and Assessment Implementation Timeline*. http://www.nysed.gov/common/nysed/files/next-gen-mathematics-instruction-assessment-timeline.pdf

New York State Education Department. (2022). *NY State Data.* https://data.nysed.gov/

New York State Office of State Assessment. (2021). *Past examinations.* https://www.nysedregents.org/

No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).

Ryan, S. V., von der Embse, N. P., Pendergast, L. L., Saeki, E., Segool, N., & Schwing, S. (2017). Leaving the teaching profession: The role of teacher stress and educational accountability policies on turnover intent. *Teaching and Teacher Education, 66,* 1-11.

Schaefer, S. (2020). *Readability, complexity, and representations: A holistic analysis of New York state algebra exams.* ProQuest Dissertations Publishing.

Senk, S. L., & Thompson, D. R. (2020). *Standards-based school mathematics curricula: What are they? What do students learn?* New York: Routledge.

Six, B., van Zimmeren, E., Popa, F., & Frison, C. (2015). Trust and social capital in the design and evolution of institutions for collective action. *International Journal of the Commons, 9*(1), 151–176.

Stanton, M., & Moore-Russo, D. (2012). Conceptualizations of slope: A look at state standards. *School Science and Mathematics, 112*(5), 270-277.

Strauss, V. (2012, May 16). *Meet the 'worst' 8th grade math teacher in NYC.* The Washington Post. http://www.washingtonpost.com/

Taylor, K. (2016, March 24). *Teachers are warned about criticizing New York State Tests.* New York Times. http://nytimes.com/

von der Embse, N. P., Pendergast, L. L., Segool, N., Saeki, E., & Ryan, S. (2016). The influence of test-based accountability policies on school climate and teacher stress across four states. *Teaching and Teacher Education, 59,* 492-502.